EDRAAK
Vol. (2025), 2025, pp. 103–110
ISSN: 3078-8412

Research Article

# Comparative Evaluation of Ensemble Learning Algorithms for Early Detection of Diabetes

Abdoul Malik [1,*],

[1] *Department of Intelligent Systems Engineering, Engineering Faculty, Ondokuz Mayıs University, Samsun, Türkiye*

**ARTICLE INFO**

Article History
Received 2 Jun 2025
Accepted: 21 Jul 2025
Revised 20 Aug 2025
Published 6 Sep 2025

Keywords
Early Diabetes Detection,
Machine learning,
Ensemble Learning,
Classification.

**ABSTRACT**

Diabetes is a severe metabolic disorder that poses a significant worldwide health challenge. Early detection is crucial to avoid serious complications and reduce healthcare burdens. Despite the rising use of machine learning in medical diagnostics, many studies rely on single algorithms, which limits predictive robustness and effectiveness. This study aims to perform a comparative evaluation of five widely used ensemble learning algorithms, including Random Forest, AdaBoost, XGBoost, LightGBM, and CatBoost on the early-stage diabetes risk prediction dataset. The algorithms have been trained and validated using 10-fold cross-validation, and their performance was assessed with accuracy, precision, recall, F1-score, and AUC. Random Forest outperformed the other models, with an accuracy of 98.46%, an F1-score of 98.75%, and an AUC of 100%, closely followed by LightGBM and CatBoost. The study's findings reveal that ensemble learning approaches, especially Random Forest, demonstrate practical and generalizable predictive performance, providing highly reliable predictions for early diabetes detection that surpass those of other methods. These findings highlight the practical value of ensemble learning in clinical contexts, offering robust, reliable, and interpretable tools to support timely diabetes diagnosis and decision-making.

## 1. INTRODUCTION

Diabetes is a chronic metabolic disorder caused by elevated blood glucose levels due to insufficient insulin production or inefficient insulin utilization by the pancreas [1]. Diabetes kills nearly 6.7 million people each year, or one every five seconds, according to the World Health Organization (WHO). It currently affects an estimated 537 million adults aged 20 to 79 worldwide, accounting for at least USD 966 billion in healthcare expenditures, a 316% increase over the previous 15 years [2]. Hyperglycemia, defined as an abnormal increase in blood glucose levels, occurs when insulin fails to transport glucose from the bloodstream to tissues, cells, and organs [3]. Diabetes, if not identified or managed, can cause serious consequences such as cardiovascular disease, nephropathy, neuropathy, retinopathy, and peripheral vascular disease[4]. Early detection is thus essential for lowering the risk of complications and improving patient outcomes. However, diagnosis is commonly delayed as a result of inadequate health awareness and a lack of hospital visits, particularly in low-income populations[5]. Clinical data and electronic medical records provide excellent resources for constructing prediction models that allow for rapid disease diagnosis and prognosis[6]. Given the growing global burden of diabetes, researchers have increasingly turned to advanced computational tools, particularly machine learning, as promising approaches for improving early detection and prediction.

Machine learning, a subset of artificial intelligence, empowers computers to execute tasks without explicit programming. In recent years, machine learning algorithms have been widely applied across diverse domains for predictive tasks, including healthcare. One of the most active study fields in machine learning is the prediction of various qualities using machine learning models on medical datasets [7]. Over recent years, there has been a growing interest in leveraging machine learning techniques for the early diagnosis of diabetes, aiming to enhance classification accuracy [8]. Scientists are particularly interested in using machine learning algorithms for medical diagnostics, and one of the most popular topics is the diagnosis of diabetes. In particular, ensemble learning has recently gained significant attention in medical diagnostics. It integrates multiple machine learning models to generate more accurate and reliable predictions than those obtained from a single model. Compared to individual classifiers, ensemble techniques improve robustness, minimize bias, reduce variation, and combine the strengths of various learners [9]. Popular methods that have demonstrated outstanding classification performance, including bagging, boosting, and stacking, are being utilized more and more in healthcare

*Corresponding author email: aabdoulmalik99@gmail.com
DOI: https://doi.org/10.70470/EDRAAK/2025/013

prediction applications. In this regard, various studies have studied the application of machine learning models for diabetes prediction, using diverse methodologies and varying levels of performance.

In this study, for predicting blood glucose levels in people with type 1 diabetes, Zhu et al. [10]suggest a fast-adaptive and confident neural network (FCNN). The recommended approach was created using three distinct data sets. The study's conclusions successfully forecasted future blood sugar levels during the 18- and 64-minute intervals. Alqushaibi et al.[11] developed a hybrid structure for Type 2 diabetes risk prediction that combines a convolutional neural network (CNN) and the Bayesian optimisation technique. Effective outcomes were achieved using the Bayesian CNN architecture. An XGBoost-based diabetes risk prediction model was proposed by Li et al.[12]. The developed method applied logistic regression, deep learning, and integrated learning techniques. The AUC (Area Under the ROC Curve) value that was found was 0.91. Tuğba [13] designed an early-stage diabetes risk prediction model using the early-stage diabetes risk prediction dataset. Traditional classifiers (Naive Bayes, J48, kNN, SMO) were compared to ensemble approaches (AdaBoost, Bagging, RSS). The AdaBoost-J48 model had the highest accuracy, and using Wrapper Subset Eval (WSE) feature selection increased accuracy and efficiency even further. Building on these findings, several studies have attempted to assess ensemble approaches and deal with data imbalance, enhancing the predictive models' resilience.

In their study, Malik et al. [14] applied a resampling technique to the Pima Indians Diabetes Database to address class imbalance. Various categorization methods have been assessed and contrasted, including Random Forest, K-Nearest Neighbour, Support Vector Machines, Artificial Neural Networks, and Decision Trees. With the IHT undersampling technique, the Random Forest reached 97.21% accuracy, 95.5% precision, 98.8% recall, 97.1% F1-score, and 99.0% AUC. In another investigation, Laila et al. [15] enhanced diabetes predictions on the early-stage diabetes risk prediction dataset by employing an ensemble learning technique. AdaBoost, Bagging, and Random Forest are among the prediction models used to compare results. Lastly, the highest accuracy of 97% was achieved by the Random Forest Ensemble Method. A diabetes prediction model based on actual data is presented by Khafaga et al. [16] For data classification, four methods were used: KNN, ANN, SVM, and DT. Compared to the other algorithms used, the results showed that KNN had the highest accuracy, at 97.36%. Moreover, other investigators have explored neural architectures and risk assessment algorithms, offering complementary perspectives on diabetes prediction.

A neural classifier-based predicting method with 82% accuracy was proposed by Pradeep et al. [17] to identify diabetes using Levenberg Marquardt back propagation (LM) algorithms with 65 hidden layers. Tan et al. [18] proposed the prediction of type 2 diabetic complications using machine learning techniques. They were successful in the random forest (RF) method's type 2 diabetic estimate as a consequence of the investigation. A Deep Density Layer Neural Network (DDLNN) has been proposed by Grupta et al. [19] for predicting diabetes. The study also included ML techniques often found in the literature, including Naive Bayes (NB), logistic regression (LR), support vector machine (SVM), K-nearest neighbour (KNN), and decision tree (DT). Consistent with the study's conclusions, it was highlighted that the suggested model outperformed alternative machine learning techniques in terms of effectiveness. Zheng et al.[20] designed a multivariate risk assessment algorithm to predict people with type 2 diabetes. The proposed model uses the LR technique based on five separate impact components. The study produced positive results, demonstrating the clinical characteristics of type 2 diabetes in depth. Despite these significant contributions, comparative reviews of cutting-edge ensemble learning algorithms remain restricted, emphasizing the necessity for systematic benchmarking to identify the most successful models for early diabetes detection using structured clinical datasets.

Although machine learning has shown potential in diabetes prediction, most research uses single algorithms or fragmented evaluations, which limit robustness and clinical usefulness. Ensemble learning provides a powerful alternative, but thorough comparisons of recent algorithms are limited. This study aims to enhance early diabetes diagnosis by developing, implementing, and systematically evaluating widely used ensemble learning algorithms to provide novel insights into their relative strengths and practical insights for their application in real-world clinical decision-support systems. Several ensemble learning techniques were used in this work to apply and assess the suggested methodology for diabetes diagnosis. Our research provides a comprehensive benchmarking of ensemble methods using structured clinical datasets. The models' performance was assessed using sensitivity, accuracy, and the Receiver Operating Characteristic curve (ROC). The proposed method, including all its steps, is explained in detail in section 2. The results are presented and evaluated in section 3, along with an in-depth discussion. Section 4 concludes with a summary of the research and suggests avenues for future research.

## 2. MATERIALS AND METHODS

For the classification tasks, we used the scikit-learn library. The model was evaluated using stratified k-fold cross-validation with $k = 10$ while maintaining the class distribution across the training and validation folds. Accuracy, precision, recall, F1-score, and AUC were among the major performance indicators used to evaluate the models in the final review phase. An overview of the suggested process is given in Fig. 1.

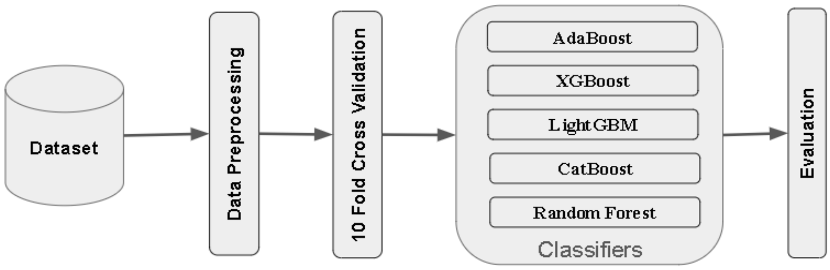


Fig. 1. Study workflow

## 2. 1 Dataset

The Early stage diabetes risk prediction dataset(UCL Database) is provided by Islam et al. [21]. Information has been gathered from Sylhet Diabetes Hospital patient records in Sylhet, Bangladesh. There are Sixteen characteristics and 520 cases associated with diabetes. One property is continuous, while the other fifteen are categorical. The data comprises 200 people without diabetes and 320 people with diabetes (Fig. 2). The collection comprises patient records from 192 females and 328 males. 90% of patients are female, and 45% of patients are male and have diabetes.

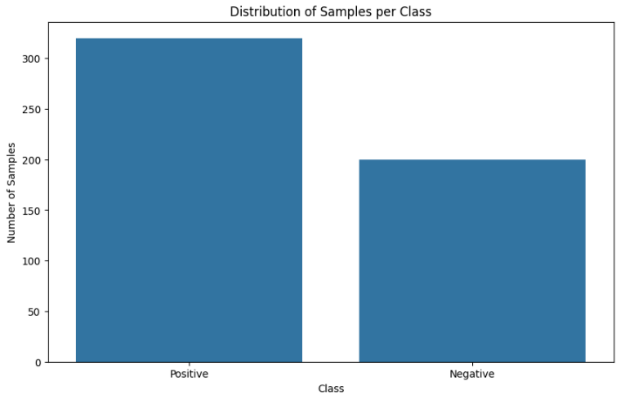


Fig. 2. Dataset distribution per class (positive and negative)

## 2. 2 Data Preprocessing

Ensuring high-quality input data is crucial for machine learning algorithms since it significantly affects the accuracy and efficiency of the model. As a result, giving careful thought to the preprocessing stage is necessary to achieve optimal performance throughout the creation of models [22]. To guarantee that random missing data do not degrade the quality of the models produced by the ML models, an appropriate imputation technique must be applied [23]. Since the dataset doesn't contain any missing values, encoding the categorical values was sufficient for preparing the dataset. In the analysis, categorical values that are 'yes' or 'positive' are represented by 1, while 'no' or 'negative' are represented by 0.

# 3. MACHINE LEARNING METHODS

Various machine learning algorithms have been developed to address different problems, among which ensemble learning has emerged as an efficient approach. Ensemble learning is a data mining technique that combines numerous techniques into a single optimal predictive model to reduce bias and variation or improve predictions [24]. This method offers superior prediction performance in comparison to a single model. In this study, the early stages of diabetes were predicted using the Random Forest, AdaBoost, XGBoost, LightGBM, and CatBoost ensemble techniques.

Random Forest is based on the bagging (Bootstrap Aggregating) technique, which builds multiple decision trees and aggregates their predictions to improve accuracy and robustness. During training, it builds a large number of decision trees and outputs the mean prediction for regression tasks or the mode of their forecasts for classification tasks. Because each tree is constructed using a random selection of characteristics and a bootstrap sample of the data, overfitting decreases and tree diversity increases. High-dimensional datasets can benefit from Random Forest's high accuracy, estimates of feature importance, and exceptional resilience to noise and outliers. It is also computationally efficient for large-scale issues due to its parallelizable structure [25].

A boosting algorithm called AdaBoost (Adaptive Boosting) turns several weak learners, usually shallow decision trees, into one powerful learner. For later models to concentrate more on challenging situations, they work stepwise by varying the weights of training samples and raising the weights of instances that are incorrectly classified. With a weight based on accuracy, each weak learner adds to the final forecast. AdaBoost is

straightforward but effective, and it often works best in circumstances with little noise and clean data. However, because it attempts to fix every inaccuracy in the dataset, it may be vulnerable to outliers and noisy data [26].
An improved and scalable gradient boosting machine (GBM) implementation is called Extreme Gradient Boosting, or XGBoost. It improves standard boosting techniques by employing shrinkage, column subsampling, regularisation (L1 and L2), and tree pruning techniques. These enhancements prevent overfitting and enhance generalisation. Cache-aware access patterns and parallel computing are further features that make XGBoost incredibly effective for big datasets. Its efficacy has been shown in real-world applications and multiple machine learning contests. XGBoost regularly produces strong prediction performance, making it appropriate for structured or tabular data [27].
Developed by Microsoft, LightGBM (Light Gradient Boosting Machine) is a distributed, quick, and powerful gradient boosting framework. It introduces key innovations such as histogram-based decision tree learning, leaf-wise tree growth strategy with depth constraints, and optimised handling of categorical features. These features allow LightGBM to achieve faster training speed, lower memory usage, and higher accuracy than traditional gradient boosting algorithms. LightGBM is especially effective on large datasets with many features and supports efficient parallel and GPU training, which makes it highly scalable for industrial-scale problems [28].
A gradient boosting approach called CatBoost is particularly effective for datasets containing categorical features. CatBoost handles categorical data natively using "ordered boosting," which avoids target leakage and overfitting, in contrast to other techniques that necessitate intensive preprocessing and one-hot encoding. It also uses sophisticated encoding methods like permutation-driven statistics and supports multiclass classification, missing value handling, and robust regularisation. CatBoost achieves state-of-the-art performance across various domains while maintaining ease of use and minimal preprocessing requirements [29]. The hyperparameters and their values for each model used in this work are shown in Table I.

TABLE. I. HYPERPARAMETER SETTINGS FOR EACH MODEL AND VALUES

| Algorithm | Hyperparameters and value |
|---|---|
| Random Forest | n_estimators=500, class_weight="balanced", criterion='gini', max_depth=32, random_state=42 |
| AdaBoost | n_estimators=500, algorithm='SAMME', random_state=42 |
| XGBoost | n_estimators=500, eval_metric='logloss', use_label_encoder=False, random_state=42 |
| LightGBM | n_estimators=500, verbose=-1, random_state=42 |
| CatBoost | n_estimators=500, verbose=0, random_state=42 |

## 4. PERFORMANCE EVALUATION METRICS

Five standard performance metrics, accuracy, precision, recall, F1-score, and area under the ROC curve (AUC), are used to evaluate a model's efficacy. Each of these metrics captures a different facet of classification quality [30].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

A True Negative (TN) indicates a correctly identified negative class, while a True Positive (TP) is a correctly identified positive class. False Negative (FN) is a missed positive case, whereas False Positive (FP) is an incorrect positive forecast for a negative sample.

## 5. RESULTS AND DISCUSSION

The Random Forest algorithm showed the best overall performance among the ensemble algorithms evaluated for early diabetes detection, with an accuracy of 0.9846, and precision, recall, and F1-score of 0.9875, along with an AUC of 1, indicating an almost perfect ability to distinguish between the two classes, with very few false positives or negatives. LightGBM followed closely behind, with an accuracy of 0.9827 and an AUC of 1, demonstrating excellent training capacity from the data. CatBoost, meanwhile, stands out with an accuracy of 0.9769, an F1-score of 0.9812 and an AUC of 1, proving its effectiveness while efficiently handling categorical variables. XGBoost also maintains solid performance with an accuracy of 0.9712, an F1-score of 0.9764 and an AUC of 0.9947, illustrating a good balance between precision and recall. In contrast, AdaBoost performs less well than the other methods, with an accuracy of 0.9269, a recall of 0.9312 and an AUC of 0.9730, suggesting that it is less robust for this specific task. The detailed results are given in Table II and Fig. 3.

TABLE. II. CLASSIFICATION PERFORMANCE OF FIVE ENSEMBLE LEARNING ALGORITHMS (RANDOM FOREST, ADABOOST, XGBOOST, LIGHTGBM, CATBOOST) ON THE EARLY-STAGE DIABETES DATASET, SHOWING ACCURACY, PRECISION, RECALL, F1-SCORE, AND AUC.

| Algoritms | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) | AUC (%) |
|---|---|---|---|---|---|
| Random Forest | 98.46 | 98.75 | 98.75 | 98.75 | 100 |
| LightGBM | 98.27 | 98.44 | 98.75 | 98.60 | 100 |
| CatBoost | 97.69 | 98.43 | 97.81 | 98.12 | 100 |
| XGBoost | 97.12 | 98.41 | 96.88 | 97.64 | 99.47 |
| AdaBoost | 92.69 | 94.90 | 93.12 | 94.01 | 97.30 |

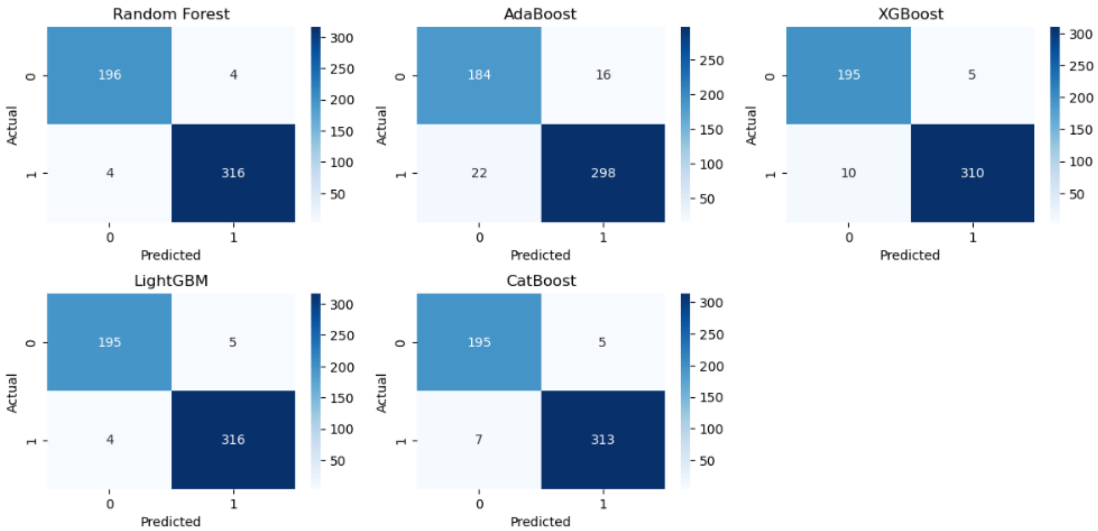


Fig. 3. Confusion matrices for the five ensemble learning algorithms used in this study. Each matrix illustrates the number of true positives, true negatives, false positives, and false negatives for early diabetes classification

As shown in Fig. 3, Random Forest, LightGBM, and CatBoost display near-perfect classification with very few misclassifications. In contrast, AdaBoost showed relatively higher errors, particularly in distinguishing diabetic from non-diabetic cases. A comparative analysis of the five models reveals that Random Forest, LightGBM, and CatBoost perform the best overall, scoring highly on every criterion. With a marginally higher AUC (AUC = 1.0) and F1-score, Random Forest appears to have stronger class discrimination and generalization. These three models closely resemble XGBoost, which has very competitive outcomes and an ideal mix between precision and recall, making it appropriate for situations where a balance between false positives and false negatives is needed. Conversely, although effective, AdaBoost exhibited comparatively lower performance. This limitation may be explained by its sensitivity to noisy data and its greater tendency to overfit on complex samples. This highlights the robustness of bagging and boosting methods in improving diagnostic accuracy. Comparison of model performance across five ensemble algorithms using multiple evaluation metrics (Accuracy, Precision, Recall, F1-score, and AUC) is shown in the Fig. 4.

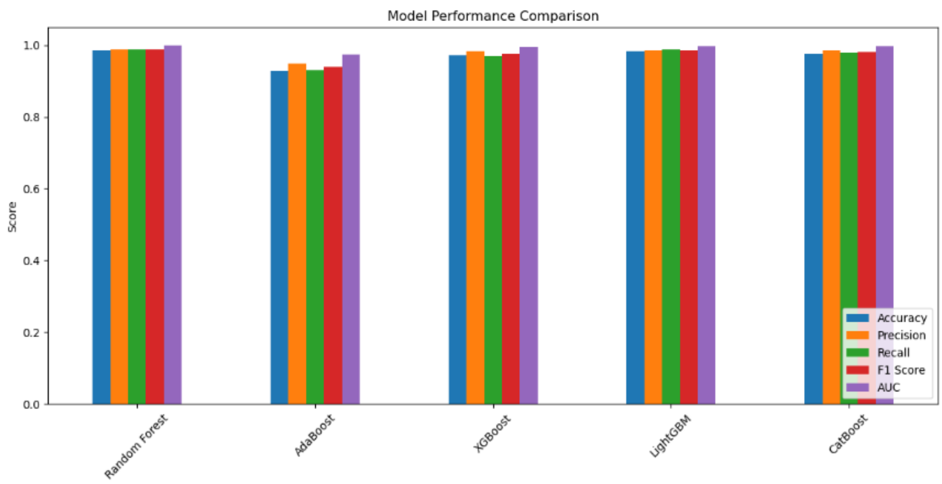


Fig. 4. Comparative performance of ensemble methods across evaluation metrics, highlighting Random Forest, LightGBM, and CatBoost as the best performers

Fig. 5 illustrates that Random Forest, LightGBM, and CatBoost achieved nearly perfect class discrimination (AUC = 1.0), underscoring their reliability for this critical diagnostic task. XGBoost also demonstrated strong performance (AUC = 0.99), reflecting its robustness and balanced predictive capacity. In contrast, while AdaBoost attained a respectable AUC of 0.97, it proved comparatively less resilient, highlighting its limitations in distinguishing between diabetic and non-diabetic cases. These findings emphasize the advantages of modern ensemble approaches in delivering accurate and dependable support for clinical decision-making.

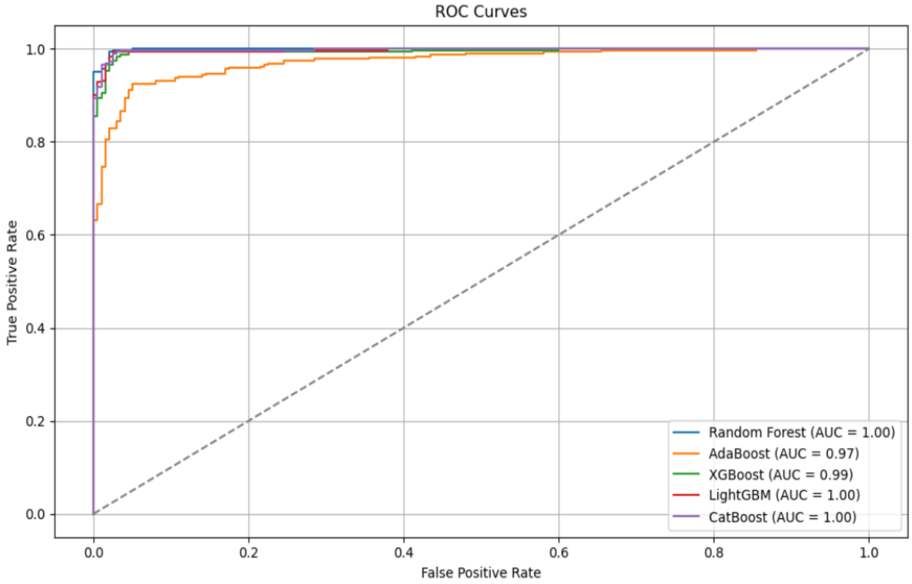


Fig. 5. ROC curves of the five ensemble algorithms with Random Forest, LightGBM, and CatBoost demonstrated near-perfect discrimination.

These results confirm the superiority of modern boosting and bagging methods, such as Random Forest, XGBoost, LightGBM, and CatBoost, in the early detection of diabetes. Among these, Random Forest is the most reliable method, thanks to its simplicity, robustness to outliers, and excellent compromise between bias and variance. In real-world scenarios where diagnostic precision is crucial, Random Forest or LightGBM are especially advised. CatBoost can be a better option when the data has many categorical variables.

A summary of the findings from this study in comparison to those from earlier studies on diabetes prediction is provided in Table III. This work's Random Forest model produced an accuracy of 98.46% and an AUC of 100%, while LightGBM and CatBoost (AUC = 1) both performed comparably well. This improvement shows that complicated patterns in clinical data can be captured using ensemble learning. The potential of Random Forest as a very dependable method for early diabetes diagnosis is highlighted. The results of this study add to the body of knowledge in academic and practice by setting a new standard for predictive performance in clinical decision-support systems. This outstanding result suggests that our method is more accurate in predicting diabetes. It therefore represents a significant advance in accuracy and efficiency in diagnosing diabetes, potentially offering clinicians a powerful tool for improving screening processes.

TABLE III. COMPARISON OF THE PROPOSED STUDY WITH EXISTING RESEARCH ON DIABETES DETECTION

| Reference | Methods | Accuracy (%) |
|---|---|---|
| Hasan Temurtas et al.[31] | MLP | 82.37 |
| Zahed and Ahmad[32] | ANN | 81.49 |
| Suyash Srivastava et al.[33] | ANN | 92% |
| Gürler et al.[34] | Random Forest | 97.30 |
| Pradeep et al.[17] | Neural Network | 96.47 |
| Laila et al.[15] | Random Forest | 97 |
| Khafaga et al[16] | KNN | 97.36 |
| Islam et al.[35] | Random Forest | 97.40 |
| This Study | Random Forest | 98.46 |

## 6. CONCLUSION

In the field of medical diagnostics, the imperative is to develop a system capable of adeptly and accurately diagnosing diseases, particularly in their early stages, consequently reducing the risk of future health challenges. In summary, this study advances the field of medical diagnostics by providing one of the most comprehensive comparative evaluations of ensemble learning algorithms for early diabetes detection. Among the tested models,

Random Forest outperformed both traditional classifiers and results reported in related literature, achieving an accuracy of 98.46% and an AUC of 100%.
These results demonstrated the effectiveness of ensemble techniques in improving clinical reliability, robustness, and diagnostic accuracy. This work is novel in that it provides actionable insights for the development of healthcare decision-support systems by establishing a systematic benchmark across several ensemble models. The results indicate that Random Forest and LightGBM can be integrated into effective screening frameworks, enabling timely interventions and improved patient outcomes. Future studies could expand on this work by examining hybrid ensemble–deep learning architectures and validating their applicability in real-time diagnostic contexts.

**Funding:**

The research was conducted without financial contributions from external funding bodies, foundations, or grants. The authors confirm that all research costs were covered independently.

**Conflicts of Interest:**

The authors declare no conflicts of interest in relation to this study.

**Acknowledgment:**

The authors acknowledge their institutions' substantial moral support and availability of research resources.

## References

[1] K. Kannadasan, D. R. Edla, and V. Kuppili, "Type 2 diabetes data classification using stacked autoencoders in deep neural networks," *Clin. Epidemiol. Glob. Health*, vol. 7, no. 4, pp. 530–535, 2019, [Online]. Available: https://doi.org/10.1016/j.cegh.2018.12.004
[2] International Diabetes Federation, *IDF Diabetes Atlas*, 10th ed., 2021. [Online]. Available: https://diabetesatlas.org/atlas/tenth-edition/
[3] K. I. Rother, "Diabetes treatment—bridging the divide," *N. Engl. J. Med.*, vol. 356, no. 15, p. 1499, 2007, [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC4152979/
[4] N. D. Wong and N. Sattar, "Cardiovascular risk in diabetes mellitus: epidemiology, assessment and prevention," *Nat. Rev. Cardiol.*, vol. 20, no. 10, pp. 685–695, 2023.
[5] P. Cihan, H. Coskun, and IEEE, "Performance Comparison of Machine Learning Models for Diabetes Prediction," in *Proc. 29th IEEE Conf. Signal Process. Commun. Appl. (SIU)*, Online, Jun. 2021, pp. 1–4, [Online]. Available: https://doi.org/10.1109/SIU53274.2021.9477824
[6] Q. Saihood and E. Sonuç, "A practical framework for early detection of diabetes using ensemble machine learning models," *Turk. J. Electr. Eng. Comput. Sci.*, vol. 31, no. 4, pp. 722–738, 2023, [Online]. Available: https://doi.org/10.55730/1300-0632.4013
[7] S. Saxena, D. Mohapatra, S. Padhee, and G. K. Sahoo, "Machine learning algorithms for diabetes detection: a comparative evaluation of performance of algorithms," *Evol. Intell.*, vol. 16, no. 2, pp. 587–603, Apr. 2023, [Online]. Available: https://doi.org/10.1007/s12065-021-00685-9
[8] A. A. Taha and S. J. Malebary, "A hybrid meta-classifier of fuzzy clustering and logistic regression for diabetes prediction," *Comput. Mater. Continua*, vol. 71, no. 3, pp. 6089–6105, 2022, [Online]. Available: https://doi.org/10.32604/cmc.2022.023848
[9] P. Mahajan, S. Uddin, F. Hajati, and M. A. Moni, "Ensemble learning for disease prediction: A review," *Healthcare*, vol. 11, no. 12, p. 1808, 2023, [Online]. Available: https://doi.org/10.3390/healthcare11121808
[10] T. Zhu, K. Li, P. Herrero, and P. Georgiou, "Personalized blood glucose prediction for type 1 diabetes using evidential deep learning and meta-learning," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 1, pp. 193–204, 2022, [Online]. Available: https://doi.org/10.1109/TBME.2022.3187703
[11] A. Alqushaibi et al., "Type 2 diabetes risk prediction using deep convolutional neural network based-bayesian optimization," *Comput. Mater. Continua*, vol. 75, no. 2, 2023.
[12] L. Li et al., "Machine learning for predicting diabetes risk in western China adults," *Diabetol. Metab. Syndr.*, vol. 15, no. 1, p. 165, 2023.
[13] T. Palabaş, "Early-stage diabetes risk prediction using machine learning techniques based on ensemble approach," *Eskişehir Tech. Univ. Sci. Technol. J.-C Life Sci. Biotechnol.*, vol. 13, no. 2, pp. 74–85, 2024.
[14] A. Malik and C. Tepe, "Improving early diabetes diagnosis: A machine learning approach for class imbalance mitigation," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, 2025.
[15] U. E. Laila, K. Mahboob, A. W. Khan, F. Khan, and W. Taekeun, "An ensemble approach to predict early-stage diabetes risk using machine learning: An empirical study," *Sensors*, vol. 22, no. 14, p. 5247, 2022.
[16] D. S. Khafaga, A. H. Alharbi, I. Mohamed, and K. M. Hosny, "An integrated classification and association rule technique for early-stage diabetes risk prediction," *Healthcare*, vol. 10, no. 10, p. 2070, 2022.
[17] J. P. M. Harikrishnan and K. Vijayakumar, "Performance analysis of neural network based classifiers for the prediction of diabetes," *Mol. Sci. Appl.*, vol. 2, pp. 24–28, 2022.
[18] K. R. Tan et al., "Evaluation of machine learning methods developed for prediction of diabetes complications: a systematic review," *J. Diabetes Sci. Technol.*, vol. 17, no. 2, pp. 474–489, 2023.
[19] N. Gupta, B. Kaushik, M. K. Imam Rahmani, and S. A. Lashari, "Performance evaluation of deep dense layer neural network for diabetes prediction," *Comput. Mater. Continua*, vol. 76, no. 1, 2023.
[20] J. Zheng et al., "Development and validation of a multivariable risk prediction model for identifying ketosis-prone type 2 diabetes," *J. Diabetes*, vol. 15, no. 9, pp. 753–764, 2023.
[21] M. Islam, R. Ferdousi, S. Rahman, and H. Y. Bushra, "Likelihood prediction of diabetes at early stage using data mining techniques," in *Computer Vision and Machine Intelligence in Medical Image Analysis*, Springer, 2020, pp. 113–125.
[22] M. O. Edeh et al., "A classification algorithm-based hybrid diabetes prediction model," *Front. Public Health*, vol. 10, p. 829519, Mar. 2022, https://doi.org/10.3389/fpubh.2022.829519
[23] J. Ramesh, R. Aburukba, and A. Sagahyroon, "A remote healthcare monitoring framework for diabetes prediction using machine learning," *Healthc. Technol. Lett.*, vol. 8, no. 3, pp. 45–57, 2021.

[24] Z.-H. Zhou, "Ensemble learning," in *Machine Learning*, Springer, 2021, pp. 181–210.
[25] A. Parmar, R. Katariya, and V. Patel, "A review on random forest: An ensemble classifier," in *Int. Conf. Intell. Data Commun. Technol. IoT*, 2018, pp. 758–763.
[26] C. Ying, M. Qi-Guang, L. Jia-Chen, and G. Lin, "Advance and prospects of AdaBoost algorithm," *Acta Autom. Sin.*, vol. 39, no. 6, pp. 745–758, 2013.
[27] J. Chen, F. Zhao, Y. Sun, and Y. Yin, "Improved XGBoost model based on genetic algorithm," *Int. J. Comput. Appl. Technol.*, vol. 62, no. 3, pp. 240–245, 2020.
[28] J. Cai, X. Li, Z. Tan, and S. Peng, "An assembly-level neutronic calculation method based on LightGBM algorithm," *Ann. Nucl. Energy*, vol. 150, p. 107871, 2021.
[29] S. Geeitha, K. Ravishankar, J. Cho, and S. V. Easwaramoorthy, "Integrating cat boost algorithm with triangulating feature importance to predict survival outcome in recurrent cervical cancer," *Sci. Rep.*, vol. 14, no. 1, p. 19828, 2024.
[30] S. Purves, B. Alaei, and D. Lolis, "Towards subsurface ML metrics," in *81st EAGE Conf. Exhib. 2019 Workshop Program*, vol. 2019, no. 1, pp. 1–5.
[31] H. Temurtas, N. Yumusak, and F. Temurtas, "A comparative study on diabetes disease diagnosis using neural networks," *Expert Syst. Appl.*, vol. 36, no. 4, pp. 8610–8615, 2009.
[32] Z. Soltani and A. Jafarian, "A new artificial neural networks approach for diagnosing diabetes disease type II," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 6, 2016.
[33] S. Srivastava, L. Sharma, V. Sharma, A. Kumar, and H. Darbari, "Prediction of diabetes using artificial neural network approach," in *Eng. Vibrat., Commun. Inf. Process.*, ICoEVCI 2018, India, 2019, pp. 679–687.
[34] N. Gürler, E. Arslan, I. Shayea, and A. Mukhamedrakhimova, "Diabetes risk prediction with different machine learning algorithms," in *Proc. 2024 IEEE 4th Int. Conf. Smart Inf. Syst. Technol. (SIST)*, 2024, pp. 403–407.
[35] M. F. Islam, R. Ferdousi, S. Rahman, and H. Y. Bushra, "Likelihood prediction of diabetes at early stage using data mining techniques," in *Computer Vision and Machine Intelligence in Medical Image Analysis*, Springer, 2019, pp. 113–125.